

2. Tutorium

„Deskriptive Statistik“

WS 16/17

Institut für Statistik, LMU

09. November 2016

Gliederung

- 1 Auswahlverfahren
- 2 Grafiken
- 3 Empir. Verteilungsfunktion

Gliederung

- 1 Auswahlverfahren
- 2 Grafiken
- 3 Empir. Verteilungsfunktion

Auswahlverfahren - Teilerhebung

Bei Teilerhebungen wird weiter unterschieden in

- Nicht zufällige Auswahlverfahren
- Zufällige Auswahlverfahren

Nicht zufällige Auswahlverfahren

- Willkürliche Auswahl
- Bewusste Auswahl
- Quotenauswahl

Zufällige Auswahlverfahren

- Einfache Zufallsstichprobe
- Geschichtete Stichprobe
- Klumpenstichprobe

Einfache Zufallsstichprobe

- Alle Auswahlmöglichkeiten haben die gleiche Wkt.

Geschichtete Stichprobe

- Uneingeschränkte Zufallsauswahl aus den einzelnen Schichten
- Alle Schichten sind (mit im Vorhinein festgelegter Anzahl) in der Stichprobe vertreten
- Effizienzgewinn, falls relevante Merkmale innerhalb der Schichten homogen sind (Elemente innerhalb der Schichten sind „ähnlich“) und zwischen den Schichten heterogen sind (Elemente aus verschiedenen Schichten sind „unterschiedlich“)
- Verschiedene Strategien zur Größe der Stichproben in den Schichten
- Gewichtung bei der Schätzung von Parametern beachten

Klumpenstichprobe

- Ziehen ohne Zurücklegen von (ganzen) Gruppen
- Verfahren geeignet, da Erhebungsaufwand reduziert
- Effizienzgewinn, falls jeder einzelne Klumpen ein gutes Abbild der Grundgesamtheit ist
(Klumpen sind in sich heterogen, untereinander homogen)

Gliederung

- 1 Auswahlverfahren
- 2 Grafiken
- 3 Empir. Verteilungsfunktion

Histogramm - Motivation

- Metrisches Merkmal: Häufigkeitsverteilung nicht von vornherein als Balkendiagramm darstellbar
(sehr viele Balken, die fast alle die Höhe $\frac{1}{n}$ hätten)
- Lösung: Merkmal in Klassen einteilen
⇒ grafische Veranschaulichung der resultierenden Häufigkeitsverteilung in Form eines Histogramms

Histogramm - Konstruktion

- Histogrammflächen sind gleich (oder proportional zu) den absoluten (h_j) bzw. relativen Häufigkeiten f_j
- Höhe des Rechtecks über der j -ten Klasse ist gleich (oder proportional zu)

$$\frac{h_j}{d_j} \text{ bzw. } \frac{f_j}{d_j}$$

- Klassenbreite $d_j = e_j - e_{j-1}$
- $e_j \hat{=}$ obere Klassengrenze des j -ten Intervalls
- $e_{j-1} \hat{=}$ untere Klassengrenze des j -ten Intervalls

Stamm-Blatt-Diagramm

- Semigrafische Darstellung für metrische Merkmale
- Beispiel:

Einheit 2		7 = 270
1		
2		0678
3		0127
4		02267
5		2688

Gliederung

- 1 Auswahlverfahren
- 2 Grafiken
- 3 Empir. Verteilungsfunktion

Empir. Verteilungsfunktion

- „Welcher Anteil der Daten ist kleiner oder gleich einem interessierenden Wert x “
- Mindestens Ordinalskala nötig
- Bilde absolut kumulierte Häufigkeitsverteilung des Merkmals X
 $H(x) = \text{Anzahl der Werte } x_i \text{ mit } x_i \leq x$
- Empir. Verteilungsfunktion
 $F(x) = \frac{H(x)}{n}$

Empir. Verteilungsfunktion - Eigenschaften

- Monoton wachsende Treppenfunktion
- Springt an Ausprägungen um entsprechende relative Häufigkeit
- Rechtsseitig stetig